DRAFT

WHITE PAPER
ON THE
BAY-DELTA PROCESS

August 23, 1995

CALFED
BAY-DELTA
PROGRAM

B-003626

The Bay-Delta Process: Logic Flow, Map to Pages

Step 1. Define the Problem 4

Problem ⟵ Objective

Cause ⟶ Action

causal linkages

Step 2. Define the Values to be Used to Evaluate Long-Term Solution Alternatives 8

Step 3. Create Solution Alternatives..... 14

Step 4. Assess Impacts, Performance of Each Solution ... 18

Step 5. Improve and Refine Solutions 23

DRAFT WHITE PAPER ON BAY-DELTA PROCESS

Table of Contents

Introduction

This "White Paper" describes the steps for the CALFED Bay-Delta Program to develop a short list of long-term solution alternatives by Spring 1996. It is primarily a planning aid for project management.

The purpose of Phase 1 of the solution finding process is to arrive at a short list of solution alternatives to carry forward into Phase 2 for environmental review and analysis. More importantly, the purpose is to arrive at that short list in a way that maximizes acceptance and durability of the long-term solution that is finally adopted. To accomplish this goal, the process should be systematic, fair, documentable, and defensible. It must involve stakeholders at an appropriate level of the process to be both fair and inclusive, and to maximize "buy-in."

One key point sets the overall "tone" of the process: We need to get the stakeholders "on board" and ensure their continued support and involvement. We can do this using four mechanisms:

1. We need to make it clear to everyone that the process is "a train leaving the station," and it is in everyone's best interest to get on board. To reach our goal within the tight timeframe, the process must be systematic, with steps that are "closed out" upon completion, and producing results that are relatively fixed. Once the problem definition is set, it will not be changed (or at least not without substantial effort). Each step will be closed out as the process continues. Consequently, early involvement of the stakeholders in this process is critical.

2. The process needs to be genuinely responsive to stakeholder input. Each stakeholder must recognize that his or her input in Task 1 is reflected in Task 2, and so on. The process must demonstrate to each stakeholder that participation is in his or her best interest.

3. The process needs to have analytic content, for two reasons:

- If the process is seen as mere discussion and flipcharts, it may appear that all we are doing is political brokering, so that any progress can be reversed by future political brokering. In this case, parties who feel underrepresented in the "brokering" may feel their chances are better if they exclude themselves from or attack the process.

- The more analytic content, the less the perception that the CALFED agencies or the contractors are biased in one direction or the other, and the less any perceived bias will be seen to matter to the outcome.

4. The process needs to be iterative, building a preliminary set of alternative solutions and refining them. The preliminary set lets everyone know that there will be real results, and roughly what those results might look like. The refinement process provides the incentive to participate over the duration of the project. It also provides an opportunity to cultivate agreement, since preliminary alternative solutions can often be improved in ways that favor most parties.

There are two other points to be made about the overall process and this White Paper. First, there is no simplified, purely analytical way to pursue this process. What will work and what will not work has to be found out empirically and through stakeholder involvement, so the process has to be adaptive. Therefore much of what is laid out in this paper is "Tools and Rules," with changeable details of implementation.

Second, this paper focuses on the process steps, not on the precise mechanisms for interacting with the public, panels, agencies, Scoping Meetings, etc. While recommendations about types of information and sources are included, most decisions along those lines are left for the Program Staff to decide. As a general rule, the more thorough and comprehensive the involvement of the stakeholders and the general public in each of the steps described here, the more complete and durable the study results will be.

The remainder of this paper describes the tasks associated with completing five steps to develop a short list of solution alternatives. The paper describes the following five steps and includes a final section covering special considerations related to the solution finding process. More detail about special elements of each step is provided in the Appendix.

Step 1 Define the Problem

Step 2 Define the Values to be Used to Evaluate Long-term Solution Alternatives

Step 3 Create Long-term Solution Alternatives

Step 4 Assess Impacts and Performance of Each Solution Alternative

Step 5 Improve Solution Alternatives

Special Considerations

Step 1. Define the Problem

Overview

The Problem Definition is important for defining the scope of the problem for ourselves, to support more effective project management. It also defines the scope to others, most importantly so that interest groups and the public can determine if they are in fact stakeholders in the decisions to be made.

The Problem Definition for the Bay-Delta Program has three parts:

- Problems to be addressed
- The geographic scope of each of those problems and the corresponding solutions
- The level of detail for data to be used in evaluating the effectiveness of possible solutions

Task Listing

The individual tasks in Step 1—Problem Definition are listed below. Descriptions of each task in the process follow.

Task 1. List the "Problems" as perceived by stakeholders and found in existing literature.

Task 2. Develop 2 to 5 "straw person" alternative solutions.

Task 3. Organize problem lists into a Problem/Objective/Cause/Action (POCA) spreadsheet.

Task 4. Organize the problems from the POCA spreadsheet into a Problems Hierarchy.

Task 5. Resolve questions (into statements of policy and intent) on geographic scope.

Task 6. Resolve questions on level of data detail.

Task 7. Synthesize Tasks 3 through 6 results into a Problem Definition.

Task Descriptions

Problem Definition Task 1

List the "Problems" as perceived by stakeholders, and found in existing literature.

This first task is to compile and list the problems related to the Bay-Delta system. Using information from previous studies, other literature, and project team experience, the Program will prepare an initial

listing to establish the basis of the problem definition statement. A first cut at this list was developed in the course of the July 5-10 preparation for the July 11 dry run, with more refined cuts following. Because the problem definition forms the basis of the solution finding process, it is critical that problems be defined in a way that highlights the important physical conditions and functions of the Bay-Delta system that stakeholders believe need to be addressed.

Problem Definition Task 2

Develop 2 to 5 "straw person" alternative solutions.

Thinking of problems in the abstract, as is done in Task 1, may miss some important problems and issues that will come up once alternative solutions are specified. A "straw person" solution alternative is an actual solution, composed of many actions. Developing "straw person" solution alternatives will help spur people's imaginations as they think of problems to be addressed. In this task, 2 to 5 "straw person" alternatives will be developed to aid the process of listing problems. These alternatives are not intended as a first cut at the preliminary alternatives, but they may serve as a way to start thinking of how best to structure alternative solutions and to lay out the set of possible alternatives.

Problem Definition Task 3

Organize problem lists into a Problem/Objective/Cause/Action (POCA) spreadsheet.

The POCA spreadsheet provides an organizing framework for much of the rest of the project, and for data management. The spreadsheet consists of four columns: Problem (stakeholder concern), Objective (underlying value to that concern), Cause (of the problem) and Action (to help meet the objective). A given problem can have multiple causes, actions, and objectives. Consequently, the spreadsheet rows are divided into "problem suites" of several rows that list all the causes, actions, and objectives associated with a given problem. At this early stage in the process, the first column (problems) of the POCA spreadsheet will be completed. Preliminary information for the other three columns will be collected and refined during Steps 2 and 3 (Values Definition and Alternatives Creation).

The POCA spreadsheet has three key roles:

- It provides a uniform format for people to list and define problems. Different people consider a problem differently. For some, a problem is "not enough salmon" (an impact-related problem). For others, a problem is "unscreened diversions" (a cause). The spreadsheet allows both sets of people to come together to contribute to the problem list. The Program and stakeholders can check the list for completeness without conflicts resulting from different ways of viewing a problem.

- It ties together the key elements of problems into a logical framework that integrates the first four sections of the project (problems, objectives, causes, and actions):

  -The Problem column will form the basis of the Problem Definition.

-The Objective column will form the basis for development of performance measures to evaluate the performance of solution alternatives (see Step 2, Task 3).

-The Action column will form the basis for listing actions to be considered for assembly into solution alternatives (see Step 3, Task 1).

-The Cause column will form the basis for developing a causal model linking potential actions to predicted outcomes (see Step 4, Task 1).

- It provides a framework for data management. A key challenge for a project of this magnitude is simply keeping track of all of the data, models, analyses and expert judgments. The POCA spreadsheet provides a clear framework for keeping track of the status of information.

Problem Definition Task 4

Organize the problems from the POCA spreadsheet into a Problem Hierarchy.

Task 4 involves developing a Problem Hierarchy to describe the levels of detail of the Bay-Delta problems. Each level of the hierarchy provides a greater level of detail and definition of the level above. The hierarchy is not intended to show the relative importance of problems, but rather show increasing level of detail defining the problems. When the hierarchy is complete, the problems will be defined in sufficient detail to guide the identification of specific actions to address the problems. A first cut at a Problem Hierarchy was developed in preparation for the July 11 dry run, with more refined cuts developed later.

The POCA spreadsheet will include indexing columns linking each problem to a particular spot or branch on the Problem Hierarchy. Using that indexing, the POCA spreadsheet and the Problem Hierarchy can be cross-checked for completeness. A key issue in defining problems on the POCA spreadsheet and the Problem Hierarchy is the phrasing of the problems such that they do not represent a narrow set of interests or dictate a specific action. The two formats (spreadsheet and hierarchy), along with a text listing of problems, taken together are a powerful, clear way to present the set of problems addressed in the project.

Problem Definition Task 5

Resolve questions (into statements of policy and intent) on geographic scope.

A critical task during the development of the Problem Definition is the delineation of the geographic scope of the problem and potential solutions. The delineation of the geographic scope will further

Value of Information Analysis

A Value of Information Analysis is an organized method to balance the extra "cost" (in time, money, and other impacts) with benefits of a broader or more thorough approach. For the Bay-Delta Program this analysis can be applied to several important decisions, such as geographic scope and data detail, to determine the most appropriate balance and provide a rationale for those decisions. For geographic scope, a Value of Information Analysis considers balancing the benefits of a larger geographic scope against the "cost" of the effort involved.

define the list of problems to be addressed. The current thinking focuses on problems within, or closely linked, to the geographic boundary of the Delta and Suisun Marsh. Problems outside the geographic scope of the program will not be addressed or will be explicitly linked to the solution finding process as inputs to or outputs from the Bay-Delta system. During this task, a "value of information" analysis (see sidebar) will assist in determining the best balance of tradeoffs between a broad geographic scope (longer, more complex, more costly analysis) and a narrow scope (shorter, simpler, less costly analysis).

Problem Definition Task 6

Resolve questions on level of data detail.

Determining the level of data detail for the analysis helps clarify the complexity of the analysis, the time and resources necessary to conduct the analysis, and the potential sources of information. During Task 6, the Program Staff will review data needs and conduct a "value of information" analysis to determine the most appropriate level of detail to develop an acceptable short list of feasible alternatives by Spring 1996. Key issues to be considered by the Program will be the type of information available for the Bay-Delta system, data gaps and information needs, the effort involved in filling data gaps, and the expectations of stakeholders regarding data analysis.

Problem Definition Task 7

Synthesize Tasks 3 through 6 results into a Problem Definition.

During this task, the results of Tasks 3 through 6 will be assembled into a Problem Definition report. The intent of this report is to clearly delineate the problems to be addressed by the project, provide a detailed description of each problem, define the geographic scope of problems and potential solutions, and describe the level of data detail for developing and evaluating potential actions and solution alternatives. This report can then be reviewed by the public and CALFED. When finalized, the Problem Definition serves as the foundation for developing and evaluating potential solution alternatives.

Step 2. Define the Values to be Used to Evaluate Long-term Solution Alternatives

Overview

The CALFED Bay-Delta solution finding process focuses on evaluating alternatives in a way that:

- is systematic, fair, documented and defensible;
- represents the underlying values of the stakeholders; and
- makes it possible to improve upon and select alternative solutions from among a large number of alternatives

The Program is facing quite a challenging evaluation problem. Each solution alternative can consist of 40 or more separate actions acting upon an inherently complex social/ecological/economic system. These actions will result in impacts on a large number of stakeholder groups, who each have a different set of underlying values. The process described here is a systematic methodology that breaks down the evaluation challenge into manageable pieces.

Three elements of the methodology are important during Step 2:

1. Because the acceptance of solution alternatives by the various stakeholder groups is critical to success, the methodology is designed to identify the values and objectives of the stakeholder groups and explicitly evaluate the performance of alternatives against them. Focusing on objectives increases the likelihood of agreement by moving stakeholders away from specific solutions, where there is often disagreement, and toward interests, where there is greater acceptance of other positions. The discussion of objectives and interests among stakeholders also encourages empathy among stakeholders and involves them in developing a common product— in this case the POCA spreadsheet and its value-related components: problem statements, objectives, and performance measures.

2. The methodology is based on a structured evaluation model to effectively compare potential solution alternatives. This model is built on performance measures that meaningfully represent stakeholder objectives and interests. This approach allows the program to evaluate potential solutions suggested by agencies, stakeholders, and the public.

3. Because selecting alternatives that best balance stakeholder interests or objectives is a great challenge for policy makers, the methodology will keep the evaluation parameters of each stakeholder group separate. The analysis will deliver multiple rankings, designed to support the CALFED Board decision to select a short list of alternative solutions to carry forward into Phase 2. Ranking the alternatives several ways, one ranking for each different set of stakeholder values,

allows the Program to take into account the diversity of values involved and evaluate alternative solutions by such factors as breadth of support and equity.

Task Listing

The individual tasks in Step 2—Values Definition are listed below. Descriptions of each task in the process follow.

Task 1. Develop overall strategy to achieve agreement on the short list of solution alternatives.

Task 2. Organize objectives into an Objectives Hierarchy.

Task 3. Develop performance measures from the Objectives Hierarchy.

Task 4. Elicit weights.

Task 5. Develop an evaluation strategy.

Task 6. Develop a Mission Statement from the Objectives Hierarchy.

Task Descriptions

Values Definition Task 1

Develop overall strategy to achieve agreement on the short list of solution alternatives.

During this task, the Program will develop the strategy and evaluation approach to foster agreement on the short list of alternatives. While the methodology described here is highly structured, in fact, there are several ways to steer the method in such a way that the chances of a broad agreement on a short list of alternatives are maximized. Key considerations that can promote agreement as the process proceeds include:

- Choices concerning how alternatives are assembled and refined to maximize breadth of support or equity
- Choices concerning how many panels of people to elicit input regarding performance measures and weights and who will participate in the panels
- Choices concerning how performance measures are selected and worded, and how technical experts and the public are involved in determining the relative weights among performance measures.

Ideally, these issues should be decided to maximize the effectiveness of five devices to aid agreement that come from the field of negotiation:

1. Develop the set of objectives and performance measures into a set that goes as far as possible toward being a "Common Language of Understanding." That is, select the performance measures in a way

that most or all of the stakeholders can agree that all measures in the set are appropriate. Try to select and word performance measures to avoid cases where one stakeholder group wants more of a measure, and another group wants less of it.

2. Seek out measures that capture "common attributes," i.e., common resources where usage patterns are such that many stakeholders jointly benefit with careful stewardship.

3. Seek out measures that emphasize the shared values of the different stakeholder groups. That is, seek out and emphasize patterns of measures and tradeoffs among measures that are similar across stakeholder groups. This is somewhat more ambitious than, but along the same lines as, the first device above.

4. Seek out cases where stakeholder groups differ in their tradeoffs in such a way that those differences allow for "win-win" movements in refining the solutions. For example, during the development of the Camp David Agreement the Egyptians said they cared most about avoiding the Israeli flag flying over the Sinai. The Israelis said they cared most about avoiding the presence of Egyptian tanks on their border. The solution: a demilitarized Sinai.

5. Seek out concepts of equity and/or fairness that are shared among the stakeholders, then use those to generate and refine alternatives. That is, develop a joint, acceptable understanding of what an equitable solution would be. There are in fact several possible ways that equity and fairness can be defined. Picking the concepts that are most apt to lead to solutions that many stakeholders can agree scores well on equity and/or fairness will foster agreement.

Values Definition Task 2

Organize objectives into an Objectives Hierarchy.

During this task the Program will develop a complete array of program objectives. The objectives define what is to be accomplished by potential solutions. Beginning with the "underlying value" information in the POCA spreadsheet, the Program will organize the information into an Objectives Hierarchy with the Mission Statement at the top. Below the Mission Statement will be listed the goals to achieve the mission, the objectives that are parts of the goals, and so on, breaking the objectives down into more detail in each step until they are specific enough to define a set of performance measures. The performance measures can then be used to evaluate alternatives (see Task 3). An example branch of an objectives hierarchy is included in the Appendix. These objectives will be based on the problems listed in the POCA spreadsheet, issues related to the implementation of potential solutions, and issues related to the impacts of potential solutions. A first cut at an Objectives Hierarchy was developed in the course of preparation for the July 11 dry run, with more refined cuts following. This first cut requires improvement based on the results of the first workshop, literature review, and other input to rearrange it into a basis for a well-organized set of performance measures.

Values Definition Task 3

Develop performance measures from the Objectives Hierarchy.

From the Objectives Hierarchy, a set of specific performance measures will be developed. A performance measure is a measurable index that assesses how well a solution addresses a stakeholder concern. The Program must develop a set of objective performance measures that meet the following tests:

- The measure must reflect a concern of at least one stakeholder, i.e., what it measures must be something that at least one stakeholder cares about.
- It must be possible to assess the measure within the reasonable cost and timeframe of the Program.
- The measure must be in units that are understandable to the participants.
- The measure must discriminate among solution alternatives. That is, the measure must result in different scores for at least one pair of alternatives. For example, global warming may be a concern of some stakeholders, but if we can't discern any difference in implications for global warming between the alternative solutions, then it should not have a performance measure.

Developing effective performance measures for the Bay Delta Program will be a challenging effort. The timeframe, resources, and data available to assess performance of alternatives are limited. The Program will have to be creative in developing meaningful, effective performance measures given these constraints.

Likewise, developing performance measures that are understandable to all participants will be challenging. Because a future step asks participants to trade off performance measures (assign relative importance), the measures must be in understandable terms and units. The Appendix includes some examples of these challenges and suggested approaches.

Values Definition Task 4

Elicit weights.

Evaluating solution alternatives requires both performance measures and weights. During this task, the Program will develop sets of weights for the performance measures, one set for each stakeholder perspective.

Developing weights for the performance measures is necessary for three reasons:

- They allow you to represent the values of different stakeholder groups in a documentable, understandable way.
- They allow you to quickly and easily rank any number of alternative solutions using any number of performance measures, by any number of sets of stakeholder weights.

- They allow you to at least semi-automate the process of determining how to change alternative solutions to improve them.

First, if there are 40 performance measures, the Program will need some way to aggregate these measures into a score for the alternative. Calculating a weighted sum of the performance measures for each alternative is a simple way to get an aggregate score to compare alternatives.

Second, the weights assigned by a stakeholder group effectively represent the tradeoffs that group would make among the performance measures. Because there would be a different set of weights elicited from each "stakeholder perspective," the ranking of alternatives for each stakeholder perspective would reflect the values and tradeoffs for that group. If we identify, say, four perspectives, we'd have four sets of weights and four rankings of alternative solutions. This approach will allow the Program to explicitly identify alternatives that receive broad support or improve conditions for each stakeholder group equally.

The method used to elicit weights is based on "multi-attribute utility analysis," a subset of operations research[1] . There are four steps in that method for eliciting weights:

1. Identify the set of performance measures that best captures ("scores") how well the alternative solutions perform with respect to the underlying values of the stakeholders. This step will be accomplished in Task 3 from the Objectives Hierarchy and a review of the feasibility of assessing those performance measures from available data, models, analyses, and expert judgment.

2. Assess or estimate how well each of the current set of solution alternatives performs on each of the performance measures. This step determines the range, from best to worst performance across the alternatives, for each performance measure. The ranges are a necessary part of the weighting process.

3. Rank the performance measures by relative weight, using graphical devices such as placards and a structured set of questions to ask of panels of people. Using recently developed calculations, that ranking can be converted to approximations of the numerical weights.

4. Where more accuracy is desired, the numerical weights can be determined more precisely using a structured "elicitation protocol" of questions to ask of the panels. For a problem as important as this Bay-Delta Project, this extra step is worthwhile. However, it does entail many hours of panel time.

The Program Staff must consider several issues related to the elicitation of weights. First, the participants and structure of the panels will need to be determined. Second, the Program Staff may be concerned about issues related to the Federal Advisory Committee Act (FACA). These panels do not function in any advisory capacity that would make them subject to FACA. A discussion of several types of panels and the associated pros and cons and the FACA issue is included in the Appendix.

---

[1] described in a book by Ralph Keeney and Howard Raiffa: Decisions With Multiple Objectives: Preferences with Value Tradeoffs (Republished by Cambridge University Press in 1993).

Values Definition Task 5

Develop an evaluation strategy.

During this task, the Program will develop the overall evaluation strategy for measuring the effectiveness and desirability of solution alternatives. This project has three roles for an evaluation strategy. In chronological order of their use, they are:

1. To provide guidance for assembling ("bundling") actions into promising alternative solutions
2. To provide guidance for refining alternatives to make them more attractive
3. To provide guidance for finally selecting the short list of alternatives.

It is important to develop an effective evaluation strategy that spans all three roles. For each role, the evaluation strategy needs to combine three types of evaluation measures:

- Weighted sum of performance measures reflecting the impacts and performance of alternative solutions (as discussed under Task 4).
- Other measures and considerations of process and institutional desirability, such as measures of impact on existing resources and measures of feasibility and implementability.
- Key measures and considerations relating to overall desirability, such as breadth of support and equity.

Discussion of each of these three areas is included in the section on "Three Types of Performance Measures" in the Appendix.

Values Definition Task 6

Develop a Mission Statement from the Objectives Hierarchy.

When the Objectives Hierarchy is complete, the Mission Statement can be developed from the information included on the Hierarchy. The Mission Statement should effectively capture the performance objectives and other evaluation measures developed in Tasks 2 and 3. The Mission Statement should reflect the two or three highest levels of objectives and can make reference to the Objectives Hierarchy itself, citing its lower branches as definitions of the terms used in the Mission Statement.

Step 3. Create Long-term Solution Alternatives

Overview

One output of the POCA spreadsheet will be a list of possible actions. From that list, actions will be assembled into several promising "packages," where each package is a solution. It will be a challenge to decide what actions can work together well to address all the various stakeholder concerns in a desirable way. Because the set of possible solutions will be refined and improved through several iterations in Steps 4 and 5, the first pass at developing alternatives does not generate perfect alternatives.

The process outlined in this step is intended to develop a list of potential solution alternatives that reflect different approaches to the objectives, without polarizing interest groups. Developing initial alternatives can be completed effectively using the measures of desirability developed in the previous Step (Values Definition Task 5). These measures are listed in the sidebar and described in the Appendix. These measures provide a way to guide the assembling of actions into alternatives.

Measures of Desirability

- Breadth of Support
- Simulated Majority Rule
- Simulated Bargaining
- Maximize Equity of Outcome
- Maximize Equity of Change from Status Quo
- Maximize Efficiency
- Maximize Balance Among Objectives
- Balanced Restoration of Resources

Each measure is not a lock-step analytical framework for assembling actions into alternatives, but rather provides general guidance on how to assemble eight different "corners of the Common Ground." Note that the approach is not intended to assemble alternatives that maximize benefit to agriculture, the environment, or urban users, since those alternatives could have a polarizing effect. Rather, this approach provides a set of alternatives that represent different ways to approach a "Common Ground."

The last task listed in this section is to develop a No-Action Alternative. However, that development should start as soon as possible, since defining the No-Action Alternative interacts importantly with the determination of the scope of the project, including the problems to be addressed and their solutions.

Task Listing

The individual tasks in Step 3—Alternatives Creation are listed below. Descriptions of each task in the process follow.

Task 1. Develop No-Action Alternative.

Task 2. Assemble action lists.

Task 3. Identify constraints and other "rules" for developing solution alternatives.

Task 4. Assemble "synergistic sets."

Task 5. Identify "solution strategies."

Task 6. Assemble preliminary alternatives.

Task Descriptions

Alternative Creation Task 1
Develop No-Action Alternative.

During this task, the no-action alternative will be developed and refined. The no-action alternative provides a benchmark for comparing the environmental effects of the various alternatives. The no-action alternative consists of projects, regulatory requirements, policies, etc. that would be in place in the absence of implementation of one of the alternatives. The tasks in developing the no-action alternative include:

- Developing screening criteria to aid in the selection of projects, regulatory requirements, and policies, etc. that should be included in the no-action alternative.
- Identifying a list of potential no-action alternative projects, regulatory requirements, policies, etc.
- Using the criteria to screen the items on the list and determine whether there is sufficient support and commitment for their continuation or implementation to be considered as part of the no-action alternative.

The no-action alternative is also an important component of identifying the problems to be addressed and possible solutions. For the CVPIA Programmatic EIS, significant effort was spent on developing and reaching agreement with stakeholders on the assumptions to be included in the no-action alternative. The no-action alternative for the CVPIA Programmatic EIS will be used as the starting point for the screening and refinement of the components. It will be modified to reflect the most recent information available.

Alternative Creation Task 2

Assemble action lists.

During this task, the Program will assemble lists of potential actions to meet the objectives identified previously in Step 2. These lists will begin with the actions identified on the POCA spreadsheet (gathered from staff, consultant team, public workshops, and other public forums). The actions from the POCA spreadsheet will need to be supplemented with at least two other types of actions not likely to appear on the spreadsheet:

- **Adaptive Management Features.** There are significant uncertainties related to information gaps and incomplete knowledge about the relationship between actions and outcomes. As a result of these uncertainties, the most attractive solutions will include some form of deliberate adaptive management. That is, the action lists should include specific mechanisms and funding to monitor the impacts and reactions to the implemented solution, and to adjust the solution according to lessons learned.

- **Externality Pricing Features.** Solutions could involve the pricing of water exports or other water uses based on non-market as well as market considerations. That pricing could be reflected in actual pricing of the water, or it could be used to derive amounts of use allowed without actually imposing the price.

An important consideration in listing actions is the level of detail in specifying an action. In some cases, it may be most appropriate to keep the level of data detail low. For example, an action might be defined as "adding 1,000 acres of salmon spawning habitat," without specifying exactly where those acres would be. Another example would be "adding a dam on the Mokelumne River," without specifying its engineering details or location, except as necessary to determine approximate storage capacity and system level impacts such as temperature, habitat, and flows.

Alternative Creation Task 3

Identify constraints and other "rules" for developing solution alternatives.

Understanding the constraints related to each action helps determine the appropriate grouping of actions into alternatives. Constraints include issues such as capital cost, capital cost per year, water, operating costs, and laws and regulations. Without a basic understanding of these constraints, it would be feasible to assemble a solution that included all actions. During this task, each action will be defined such that the basic constraints are understood.

Key issues for the CALFED Program are related to available funding and institutional constraints. Preliminary work on the anticipated range of funding available will be important before assembling actions. Some of the best solutions might require changes in institutions, laws, or regulations. CALFED guidance on the range of institutional flexibility will also be important.

Alternative Creation Task 4
Assemble "synergistic sets."

Assembling actions into solution alternatives will be a very challenging task. During this task, assembling actions into groupings will help make assembling solution alternatives more manageable. One way to do this is to identify subsets of actions that form "synergistic sets," i.e., where there are complementary physical interactions between actions. One example would be the construction of a dam and accompanying adjustments in habitat development and management. These groupings of actions can then be assembled into alternatives in Task 6, described below.

Alternative Creation Task 5
Identify "solution strategies."

In this task, "solution strategies" will be developed to guide the assembling of alternatives. As described in the overview of this step, one way to develop solution strategies that support consensus and agreement is to base the strategies on the eight measures of desirability developed in Step 2, Task 5. Each of the eight measures would be developed into a framework, or at least very general guidance, for assembling actions into preliminary alternatives that score well on that measure.

Alternative Creation Task 6
Assemble preliminary alternatives.

During this task the results of Tasks 2, 3, and 4 are assembled into preliminary solution alternatives based on the solution strategies developed in Task 5. The key goal of this task is to develop a wide range of possible solutions that capture the objectives of the program. Because they will be improved and refined in the Steps 4 and 5, the preliminary alternatives do not need to include the actual alternatives that will be on the short list. That is, the Program will not be simply eliminating some alternatives from further consideration (screening alternatives), but rather improving, refining, and combining them to reach a shorter list.

While it is possible to develop some computational aids for performing this task, it may be that the most effective way to assemble alternatives would be to gather a small group of knowledgeable people with a large table and 3 x 5 cards describing actions and synergistic sets (with several copies of cards for each action, so an action can be included in several alternative solutions). This group could develop, over several days, several solution alternatives for each of the solution strategies. This task should be as systematic as possible to ensure that it can be documented and defended and to ensure that a wide range of possible solutions is developed.

Step 4. Assess Impacts and Performance of Each Solution Alternative

Overview

Perhaps the central analytic challenge of this process is to assess the impacts and performance of each solution alternative. Each solution may be comprised of 40 or more separate actions. Those actions act on and interact with a complex societal/ecological/economic system with large uncertainties. The resulting outcome can be described by the performance measures developed during Step 2 (possibly as many as 40). For each alternative, the performance measures must then be weighted and summed into a set of scores and rankings, one for each stakeholder perspective. This step develops an assessment approach that integrates the results of previous steps and produces an objective measurement of the performance of each alternative. During this step, an "actions-to-outcomes" model will be developed. The model will take as inputs the 40 or more separate actions comprising a solution alternative and deliver as outputs the set of performance measures that measure the performance/impacts of the alternative.

The "actions-to-outcomes model" must account for the significant linkages between actions. The model will be based on existing databases, model results, analyses, and expert judgment. While the model will contain many approximations, it will be designed to make the best use of available information to support the process of evaluating, improving and selecting the solution alternatives. The intent is not to develop a comprehensive prediction model, but rather to reflect, as accurately as possible, the causes and linkages necessary to compare performance of the alternatives.

There are several key issues related to developing the actions-to-outcomes model and measuring performance:

- **Decision-focused, iterative approach.** A key strategy for dealing with the assessment challenges will be to follow a decision-focused, iterative approach. Analyses will not be designed to gain a full understanding of the underlying processes, but will be trimmed to focus only on what is called for to rank the alternative solutions. The overall model development will be iterative, starting with a very approximate first pass, which will then be used to identify the critical variables to be modeled more carefully in the second pass and subsequent iterations.

- **High uncertainty/risk.** There is a great deal of uncertainty involved between actions and outcomes in the Bay-Delta system. That uncertainty lies both intrinsically in the societal/ecological/economic system, and in the limitations of knowledge about the system. In the model, uncertainty can be addressed explicitly by incorporating error bars to capture a range of disagreement among experts or uncertainty due to incomplete understanding of the system.

- **Adaptive Management.** Adaptive management is a second way to deal with high uncertainty, however, it presents a performance assessment challenge. It will be difficult to anticipate the long-

term performance of solutions that feature adaptive management, though it is important to give appropriate credit to solutions for their adaptive features.

- **Disagreements about data and models.** This project is marked by several different sources of data, models, analyses, and expert judgment— in some cases with important disagreements between sources. A "Second Table Process" will be a useful forum for resolving these issues. A Second Table Process would be a separate process where technical experts reach agreement on the data, models, analyses and expert judgments to use as data sources, when to span disagreements with error bars, and the process to for source accreditation.

- **Data gaps.** There is often a need for data that is not available directly from existing databases, models or analyses. In these cases, expertly elicited expert judgment can fill those gaps. There are particular expert judgment elicitation protocols designed to get the most reliable, repeatable data from an expert-judgment source, typically in the form of probability distributions. These probability distributions fit naturally into the generally probabilistic approach described here. Important considerations include the selection of the experts and the elicitation and documentation of the expert judgment.

Task Listing

The individual tasks in Step 4—Performance Assessment are listed below. Descriptions of each task in the process follow.

Task 1. Organize causal relationships into a causal model.

Task 2. Develop linkages between the causal model and actions.

Task 3. Develop linkages between the causal model and performance measures.

Task 4. Develop workplan for modeling, data collection, and expert judgment elicitation.

Task 5. Execute modeling workplan.

Task 6. Execute data collection workplan.

Task 7. Execute expert judgment elicitation workplan.

Task 8. Assemble actions-to-outcomes evaluation model.

Task 9. Assess impacts and performance of each solution.

Task Descriptions

Performance Assessment Task 1
Organize causal relationships into a causal model.

During this task, the causes identified during development of the POCA spreadsheet will be organized into a causal model. The model will show the causal relationships among components of the Bay-Delta system. Another way of describing the causal model is as an influence diagram. The causes identified during development of the POCA spreadsheet can be indexed to all of the problems identified and to other causes. This model will begin simply in a graphic form, useful for building insight into the causal relationships between the different elements of the system. The result of this task will be a graphic depiction of how elements of the system interrelate.

Performance Assessment Task 2
Develop linkages between the causal model and actions.

This task will build on the causal model by defining the linkages between causes and actions. Since the ultimate goal of this step is to produce a model that assesses the performance of each solution, the evaluation model must link the actions comprising each solution alternative (and the interactions among those actions within the Bay-Delta system) to the performance of the solution. That is, evaluating the performance of solution alternatives developed in Step 3 using performance measures developed in Step 2.

As a first step in building these linkages, this task will identify how each of the actions from Step 3 relates to the causal model developed in the previous task. These linkages will be specifically identified to build toward an evaluation model. Many of the relationships between actions and causes will be found on the POCA spreadsheet. Linkages from the POCA spreadsheet will be reviewed and supplemented as necessary by other linkages between causes and actions, such as adaptive management and water pricing.

Performance Assessment Task 3
Develop linkages between the causal model and performance measures.

The next step toward the evaluation model is to add the relationships of the performance measures to the causal model. The performance measures developed in Step 2 will be linked to the causal model by identifying the relationships between elements of the Bay-Delta system and the performance measures. The POCA spreadsheet will include many of these linkages as described by the relationships between the listed causes and the objectives developed.

Additional linkages may be called for concerning the associated impacts of potential solutions (e.g., minimize recreation impacts). Other linkages that could be called for include:

1. Indirect economic implications, including effects on local and state economies.
2. Market economic implications, including cost and production level changes that may propagate through the several markets involved.
3. Technical user reaction implications, including most notably crop substitution, but also including any other technical user reactions that may be identified.

These linkages will not be represented on the POCA spreadsheet because it focuses on the problems to be addressed and related causes, objectives, and actions.

The challenges of this task highlight some of the key issues discussed in Step 2. The performance measures must be understandable to all of the participants. The Program must develop (with the participants) a clear definition of the appropriate balance between the desired performance measures and the availability of data.

At the conclusion of this task, the Program will have the framework of the evaluation model. Task 1 developed the relationship among elements of the Bay-Delta system (how the system works). Task 2 linked the actions to the system (how actions affect certain elements of the system). Task 3 linked the performance measures to the system (how outcomes will be measured).

Performance Assessment Task 4

Develop workplan for modeling, data collection, and expert judgment elicitation.

During this task, the Program will identify the best sources of information to make the evaluation model work. Each linkage in the model is essentially a data requirement. Each linkage must be reviewed to determine if the requirement is best satisfied by existing data, models, analyses, or expert judgment elicitation. The Program will then develop a workplan for data collection, model runs, analysis, or expert judgment elicitation based on the best sources of information.

If expert judgment is called for, expert panels must be appointed. If data, model runs, or analyses are called for, the collection of this information should be planned. If there is not clear consensus on the appropriate data, model runs or analyses to use, then some sort of data, modeling, or analysis panel must be appointed to determine the best way to proceed. The panel can either decide to adopt one particular dataset/model/analysis, span alternative ones with an error-bar approach, or develop another way to combine the conflicting inputs into a consistent basis for fulfilling the data requirement. As with the weight-elicitation panels, these panels are not subject to the Federal Advisory Committee Act (FACA). They are used as information sources only. They are never asked for advice concerning the alternative solutions or the overall decision process.

Performance Assessment Task 5
Execute modeling workplan.

Performance Assessment Task 6
Execute data collection workplan.

Performance Assessment Task 7
Execute expert judgment elicitation workplan.

During these three tasks, the Program will implement the workplans for collecting the necessary information to make the evaluation model run. Each of these tasks will be well defined by the structure and content of the evaluation model and the decisions made during Task 4. Each of these tasks, however, could represent a large amount of work, time and expense, even if the time constraints of Phase 1 dictate using only existing data, models, analyses, and expert judgments in Phase 1.

Performance Assessment Task 8
Assemble actions-to-outcomes evaluation model.

During this task the products of the preceding seven tasks are assembled into a very approximate, actions-to-outcomes evaluation model. This model will be used in Task 9 to evaluate the performance of each alternative, using the weights elicited during Step 2. The format of this model is intended to be as open and accessible as possible given the complexity of the system. The model can be designed such that assumptions, linkages, and evaluation can be demonstrated and explained to participants. Further discussion of the structure of the model is included in the Appendix.

Performance Assessment Task 9
Assess impacts and performance of each solution.

In this task, the performance of each preliminary solution alternative will be evaluated using the actions-to-outcomes model developed and refined in Task 8. Each alternative will be evaluated using the sets of performance measure weights identified for each interest group in Step 2. The results will be one score representing overall desirability for each alternative for each set of weights. These results will form the basis of refining and improving alternatives in the next step.

The model will be developed iteratively, concurrently refining the model as solution alternatives are refined. At any given time, the Program will use the most refined version of the model currently available.

Step 5. Improve and Refine Solutions

Overview

Because it is unlikely that the optimal solution alternatives will be developed in the first attempt, it will be important to refine and improve alternatives several times to reach a short list. The evaluation model provides a straightforward, analytical way to demonstrate performance and identify possible refinements to improve performance. The evaluation process (Step 4, Task 9) will produce results that will demonstrate on which performance measures an alternative did not score well. With this information, refinements can be tested for an alternative. Likewise, the performance evaluation may show that two alternatives are identical in most respects, but one outperforms the other on one or two performance measures. This information allows the Program to eliminate one of the alternatives from further consideration.

In this Step the Program will review and refine alternatives to identify the most promising solution alternatives to select for Phase 2. After assembling alternatives (Step 3— Alternatives Creation) and evaluating them (Step 4— Performance Assessment), the Program will use details of that evaluation to identify ways to improve the alternative solutions, loop back through the evaluation model (Performance Assessment Step 9) with the revised alternatives, and keep revising alternatives and running them through the performance assessment model until the alternatives do not improve any more, or until the budget or time run out, whichever comes first. This is a systematic effort to come up with alternative solutions that between them represent, as well as possible, the "Common Ground" among the different stakeholders.

This iterative approach has the following benefits:

- It allows the Program to use existing data to focus further data collection, modeling, analysis and expert judgment elicitation on evaluating, ranking, and refining only those solution alternatives identified as promising. This should lead to an efficient use of analysis resources.
- It allows the Program to demonstrate responsiveness as issues and concerns are raised in each task, then incorporated in a fairly short time into the next round of runs. This continuous responsiveness should lead to a development of trust, buy-in, and constructive participation.
- It allows the Program to build up a series of small agreements about alternatives through the explicit demonstration of performance. All participants can see how well alternatives perform for each measure and where they might be improved. The net result is that the series of small agreements builds negotiation momentum toward overall agreement on a short list.
- It allows the Program to improve alternatives in ways that can be shown to benefit most (or many) participants. This approach encourages constructive participation, much more so than first

presenting highly refined alternatives, where any change favoring one stakeholder is apt to penalize another stakeholder.

Task Listing

The individual tasks in Step 5—Solution Refinement are listed below. Descriptions of each task in the process follow.

Task 1. Identify changes in alternatives to improve overall desirability of alternatives.

Task 2. Refine alternatives and re-evaluate.

Task 3. Repeat Solution Refinement Tasks 1 and 2 as necessary to achieve agreement on a short list.

Task Descriptions

Solution Refinement Task 1

Identify changes in alternatives to improve overall desirability of alternatives.

In this task, the Program will analyze the results of Step 4, Task 9 to identify changes in alternative solutions to improve overall desirability of alternatives. The weighting information from Step 2 can be used to calculate which changes (i.e., which actions revised in which direction) will lead to favorable overall changes in the performance of a solution. For example, the evaluation model will allow for an analysis to determine if refinements can be made to an alternative that performed well for most stakeholder groups to improve its performance for other stakeholder groups. At the same time, a common-sense examination of how each action affects the measures will likely also identify favorable changes.

Solution Refinement Task 2

Refine alternatives and re-evaluate.

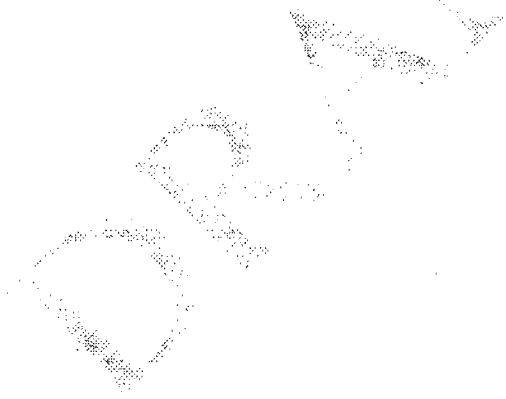
During this task, refinements identified in Task 1 will be incorporated into the alternatives to generate new alternatives. Then, by repeating Performance Assessment Task 9 the performance of each refined alternative will be evaluated.

Solution Refinement Task 3

Repeat Solution Refinement Tasks 1 and 2 as necessary to achieve agreement on a short list.

This task provides the opportunity for continued involvement of stakeholders and others in identifying improvements to alternatives. This involvement is likely to create a forum for effective discussion of the alternatives that meet the diverse needs of the stakeholders, as well as ways to make alternatives more acceptable to all interests.

An additional activity in this task will improve the validity of the evaluation and, at the same time, promote constructive stakeholder participation. As the solution alternatives are refined, different performance measures rise to the top as discriminators among the alternatives. As the importance ranking of the performance measures shifts, revisiting the weighting elicitation with stakeholders will refine the weights for measures that have newly appeared near the top of the importance ranking. This process may support more effective, ongoing, constructive participation than a relatively static process where the Program elicits weights once and then uses those weights for the rest of the project.



Special Considerations

This section describes some special considerations for the Program as the process proceeds through the five steps to develop a short list of alternatives.

Division of Labor

One of the advantages of the method presented here is that it uses the structures of decision analysis to divide up the inputs from stakeholders, technical experts, and process specialists, so that no source is asked to provide inputs they are not qualified to provide. Stakeholders are asked for their values, in terms of objectives, performance measures, and tradeoffs. Technical experts are asked for technical information and expert judgment where needed. The process and evaluation model development are guided by the process, or method, specialists. Thus, technical experts are not asked for value tradeoffs that are best performed by policy makers and stakeholders. Likewise, stakeholders are not asked to be technically competent or to evaluate alternatives; information is presented in terms that stakeholders can understand.

Openness to Suggestions for Alternatives From Other Parties

A second advantage to this approach is that it allows for evaluation of alternatives from outside the CALFED Program. Extending an invitation to any party to suggest alternative solutions would demonstrate openness and may generate new and better ideas. An invitation should include a set of specifications for how the suggested alternative is to be described, so that the Program can determine how to evaluate it. Such an invitation would have the effect of placing the Program in a neutral position, ready and willing to evaluate any alternative that is presented. If a submitted alternative scores better than Program alternatives, it can be incorporated into the list.

Democracy of Analysis

The evaluation model is intended to be understandable to most (if not all) participants. It can be constructed in a spreadsheet format (see the Appendix) so that it is accessible on most desktop computers. The Program could distribute the model and allow stakeholders and others to use the model to search for effective solution alternatives. This approach would set a very constructive tone for the process and demonstrate openness.

Possible Problems with Explicitness

There are potential concerns associated with an approach that is based on open, explicit evaluation of objectives, value tradeoffs, and outcomes. Generally, it is not in a bargainer's best interest to disclose

values and value tradeoffs. It is only in the stakeholder's best interests in the Bay-Delta process, because of the way the process is structured. But stakeholders may not realize this, and so withhold or misrepresent their value information. It could also be the case that being explicit about objectives and tradeoffs in the Bay-Delta process could reveal information that could be used to a stakeholder's disadvantage in some other bargaining process. One way to work around this problem is to elicit values and value tradeoffs from people who could be considered to know the values and tradeoffs of Stakeholder Group "A," without actually being members of that group, and without privileged information on that group. While that is less direct stakeholder representation, it does get around the disclosure problem.

Another problem of explicitness has to do with the ability of the Program to make clear the interrelationships between stakeholder positions and the water allocation situation. The analysis may have the effect of making it embarrassingly clear that certain parties enjoy large benefits relative to other parties, and that those same parties constitute a large part of the water allocation "problem." In this case, such clarity could lead to unfortunate political implications for the project. The Program should consider these issues carefully and consider measures to take, consistent with an analytically and ethically sound analysis, to make sure the Program will be able to complete the project without interruption.

Decision Quality Criteria

Whatever decisions CALFED makes, there will probably be legal actions against the result. Therefore it is paramount that the Program adhere to ironclad decision quality criteria. Four criteria have already been discussed:

1. **Systematic.** The process systematically accounts for all feasible actions, impacts, and stakeholders. Every decision follows from specified inputs through a specified process. Decisions are not arbitrary.

2. **Fair.** There is no systematic bias among actions or stakeholders.

3. **Documented.** Every decision is explained in writing that specifies its inputs and decision process. The writing is clear and accessible.

4. **Defensible.** The process is defensible, as a logical result of the previous criteria. The analysis must be formally correct. Perhaps the biggest pitfalls here are incorrect handling of weights and incomplete handling of interactions.

With due diligence, conformance with these four criteria should follow naturally from careful adherence to the process laid out in this paper.

The following is another set of criteria currently used to define "Decision Quality" by Strategic Decisions Group, one of the three leading decision analysis consulting firms:

1. **Values.** Are the values that govern the decision thoroughly sought out and clearly addressed? The Value Definition step should have the process perform well on this criterion.

2. **Frame.** Is the decision problem properly framed? The Problem Definition step should have the process perform well on this criterion.

3. **Alternatives.** Are the alternatives created in a process that systematically explores all feasible avenues to achieving the objectives? Creativity and imagination during the Alternatives Creation step will be called for to perform well on this criterion. Perhaps special attention should be paid to creative options such as institutional, legal, and pricing strategies.

4. **Information.** Is information used the best available, subject to time and budget constraints? Is it selected and managed in a decision-focused way? Is there a systematic accreditation process? This criterion is a significant challenge for the Program, because no matter how thorough the analysis, some parties will argue about the information used. Issues in this arena can be minimized by using thorough sensitivity analyses, use of error bars and probability distributions to be explicit about uncertainty, and developing alternatives that are attractive over a wide range of values of the key uncertain parameters.

5. **Logic.** Does the analysis follow the logical principles of decision analysis? The approach described in this paper should do well on this criterion.

6. **Commitment to Action** Does the process result in a situation where every person whose participation is necessary for implementation is in fact committed to that participation? This may be the most important and difficult criterion for success. Specific challenges for the Program include:

- Fostering creativity regarding institutional arrangements
- Determining which panels and groups of people are convened to be asked which questions
- Getting the right people into the room in the right frame of mind to elicit the information that will make the analysis sound.
- Ensuring that all stakeholders for each problem are represented

Appendix

This Appendix includes more detailed discussion of several topics presented in the white paper. The Appendix is organized as described below:

Example Branch of an Objectives Hierarchy

The following is an example branch of an Objectives Hierarchy. Beginning with the overall program mission, the hierarchy moves through levels of detail to a set of performance measures.

Mission Develop an integrated, long-lasting plan and implementation strategy to improve the natural environment and reliably meet the needs of the human communities that rely on the Bay-Delta System. *One way to do that is to:*

Goal Enhance the quality of the Bay-Delta ecosystem so that it supports quality habitats for varied and valuable species. *One way to do that is to:*

Objective Enhance Bay-Delta habitat. *One way to do that is to:*

Sub-Objective Enhance Bay-Delta aquatic habitat. *One way to do that is to:*

Sub-Objective Increase productivity of Bay-Delta aquatic habitat, *as measured by:*

Performance Measure Upper trophic biomass, measured on a scale made understandable to a non-technical stakeholder. (*The same sub-objective could have other performance measures, such as:)*

Performance Measure Biomass diversity, measured on a scale understandable to a non-technical stakeholder. This measure recognizes that tons of biomass alone may not be an adequate measure of habitat productivity.

Key Issues Related to Developing Performance Measures

There are four key issues related to performance measures for the Program to consider and resolve:

- **Reasonably available measurements.** One of the key issues for the Program is developing Performance Measures that are reasonably measured with available time and resources. Typically, performance measures that most directly assess how well an alternative performs with respect to an objective are not feasibly assessable from existing data, models, analyses or expert judgment. For example, a desirable performance measure could be "number of adult winter-run salmon." However, predicting that number for any given set of actions is extremely uncertain, and subject to a great deal of debate. A compromise has to be worked out between desirable performance measures and what is feasibly available from existing data, models, analyses or expert judgment. In this example, that compromise might be acres of salmon habitat of a particular quality or better, or it could be a subjective scale, as discussed below.

- **Understandable to participants.** For many aspects of impact or performance some technical measure would be the most direct measurement. For example, lower trophic productivity could be most directly measured by tons of lower-trophic biomass per acre. But if that measure is to be traded off against another measure by a generalist, then the Program must develop a scale that is more meaningful to a generalist. Is ten tons per acre a large amount or a small amount? What is the significance of six tons per acre as opposed to ten tons per acre? Why does it matter? In many cases, the best way to handle this issue is to build a subjective scale, even though a direct, numerical scale (such as tons per acres) is available.

- **Building subjective scales.** In many cases the best performance measure will be a subjective scale. For example, "habitat diversity" could be represented by a 3-, 5- or 7-level scale. Each point on this scale should be briefly defined (from one to five sentences) in such a way that two different people will understand what each point on the scale means. While it may take a specialist to develop a subjective scale, the wording of the sentences describing each point should be developed to be clear to whatever panelists are going to be asked to provide input on the relative importance of the measures.

- **Accounting for performance scales that do not linearly reflect underlying values.** In many cases, a direct numerical performance measure will not be linearly related to the underlying value. That is the value of a performance measure is not directly proportional to the quantity measures. For example, values concerning number of salmon could involve a lower threshold, below which the value drops off precipitously due to risk of extinction, and perhaps an upper threshold, beyond which extra numbers of salmon do not matter very much. More generally, many measures may have "diminishing marginal returns" — that is, the more there is of a resource, the less you care about an additional increment. Specialists, or in some cases generalists, can be asked a structured set of questions to elicit these nonlinear relationships so that the evaluation accurately reflects these types of situations.

Assigning Weights to Performance Measures

There are three important reasons that assigning weights to performance measures improves the evaluation of solution alternatives:

- **A weighted sum for each alternative simplifies the comparison of alternatives.** Comparing solution alternatives on 40 or even 10 different performance measures is a complex task. For example, comparing 8 alternatives on the basis of 10 performance measures could be presented as 8 rows (each row is an alternative) of 10 numbers each (an 8 x 10 matrix). Unless one alternative (row) is better than all the rest on all 10 performance measures, it is extremely difficult to compare and rank the solution alternatives by looking at 8 sets of 10 numbers each. There is a wealth of experimental data showing that people cannot compare and rank alternatives consistently. Calculating a weighted sum of each performance measure produces a simple-number score for each alternative, simplifying the comparison.

- **Weights incorporate an analytical rigor to the process.** If the comparison is completed without weighted sums, a good mathematician can either (1) figure out the weights (at least bounds on the weights) that were effectively used; or (2) show that your choices were not consistent with any one set of weights, in which case the process can be accused of being arbitrary, or of evaluating the alternative solutions with some unannounced criteria.

- **Scoring alternatives without weights will be arbitrary.** Constructing the evaluation without weights (summing the performance measures without weights) is equivalent to equal weighting, which is completely arbitrary. Since the weights are a function of the range of impact (or benefit) for each performance measure (from best to worst among the solution alternatives), it would only be by coincidence for there to be a set of measures and ranges for which equal weights would be appropriate. Typically, correctly identified weights for eight or more performance measures vary by a factor of at least five to one.

Evaluation using performance weights is an important component of two approaches for considering resource management decisions around the country.

1. **Externalities.** Many utilities in the Northeast have been asked by Public Utility Commissions to take into account externalities in their decision making. In response, the utilities are evaluating multidimensional outcomes of decisions by defining performance measures that capture non-economic, non-market impacts and performance, such as ecological and societal impacts. They then determine appropriate weights for those measures and select alternative solutions based upon the weighted sum of those measures. In fact, the underlying method of the approach described here, multi-attribute utility analysis, is uniquely suited to evaluation of externalities.

2. **Contingent Valuation.** This is a name given to the use of survey techniques to "price out" non-economic, non-market impacts. Although there is debate about the validity of eliciting pricing parameters from surveys, the general principles remain the same as those the Program is discussing

for this project. And again, the underlying method of our approach, multi-attribute utility analysis, is uniquely suited to "pricing out" non-economic, non-market impacts.

These two current discussions are of special interest to the Bay-Delta Program, because the Program may consider solutions that involve pricing out water exports according to their "multidimensional cost" including non-economic/non-market/ecological/societal costs. This pricing out may be reflected in actual costs charged for exports, or it may be used as a basis for determining limits on levels of exports without actually being charged. In either case, the concept of weighted sums of performance measures has direct use in the form of these important policy tools to consider in fashioning alternative solutions.

Panels for Eliciting Performance Measure Weights

The Program must consider two important questions related to assigning weights to performance measures: how the weights will be elicited and from whom. Weights can be elicited from panels of stakeholders using a placard-ranking process and/or a series of weight-elicitation questions.

Trading off performance measures. Placard-ranking or weight-elicitation questions force the respondent to "trade off" the various conflicting objectives by representing them as tradeoffs between performance measures. The respondent should never be asked to answer an elicitation question involving information he or she does not understand. Therefore, the Program should "nest" the performance measures. For example, detailed, technical measures about different habitat types will be rolled up into "equivalent wetland acres" using tradeoffs elicited from habitat specialists. Then equivalent wetland acres can be traded off against water supply measures by more generalist, policy-oriented respondents. A review of the performance measures to determine which ones require tradeoff judgments (and "nesting") from specialists, and which require judgments from generalists, will be required.

Stakeholder panels. There are several possible panels that can be used as "value sources." The decisions on which panels to use are largely policy decisions to be made by the Program. These decisions about the types of panels (and what participants) are centrally important, because they directly influence the support for the evaluation process and the credibility of the results. The following are some alternative panel types, and their pros and cons. Any combination of these panels can be used, with particular panel types for particular tradeoffs.

Direct Stakeholder Representatives.

Pros: Direct stakeholder involvement promotes stakeholder participation and fosters "buy-in." The resulting weights are the most direct measure of stakeholder values.

Cons: Some stakeholders may misunderstand the process, and so intentionally misrepresent their preferences, in spite of efforts to make it clear that it is in their best interest to be honest about their preferences. Possible undesirable turnover in panelists due to time commitment. Possible delays or need for special panels because some stakeholder representatives may feel that they must consult with constituents before answering. Challenge of assembling panels that are seen to be inclusive and balanced.

Authoritative Panel on Social Value Tradeoffs.

Pros: More impartial and less polarized perspectives. More direct savvy regarding ways in which different stakeholder perspectives can be handled. Less apt to erroneously misrepresent preferences.

Cons: Less direct stakeholder involvement, with resulting decrease in participation, direct measurement, and buy-in. Participants often have time commitment problems.

Agency Panels.

Pros: Direct representation of those organizations from whom we must get buy in. Often valuable "political overview" perspective.

Cons: Less direct stakeholder involvement, with resulting decrease in participation, direct measurement, and buy-in. Time commitment problems, since agencies may not appreciate the importance of budgeting adequate time for the participants. May be apt to erroneously misrepresent preferences. May be preoccupied with the appearance of particular tradeoff judgments, as opposed to accurately representing actual preferences.

BDAC.

Pros: Has direct validity from its appointed role. Otherwise, all the pros of an "authoritative panel."

Cons: A value source role may not be compatible with its other roles in the process. Otherwise, all the cons of an "authoritative panel."

Internal Panel.

Pros: Essentially no problems with erroneous misrepresentation. Participants are directly motivated to be impartial and to make responses that promote the political feasibility of the evaluations that result. Direct control of time commitment problems.

Cons: Less direct stakeholder involvement, with resulting decrease in participation, direct measurement, and buy-in. Limited to in-house perspectives.

Consultant Panel.

Pros: Same as for internal panel, though less direct motivation to be impartial and to promote the political feasibility of the evaluations that result. Also, if the consultants are not normally part of the project, they can be external information sources. Can pick the best expertise available on particular subjects.

Cons: Perhaps the poorest of the panel types for perceived fairness and directness of stakeholder representation, with resulting decrease in participation, direct measurement, and buy-in.

Technical Panel.

Pros: Valuable or necessary for particular tradeoffs that require special expertise, such as tradeoffs between different measures of habitat quality. Especially valuable for special-expertise tradeoffs where there are disagreements within the technical community.

Cons: Can get into problems when some stakeholders feel that certain opinions should be used that members of the specialist community feel are not technically valid. Although in these cases, convening a technical panel is almost certainly better than simply making in-house judgments.

These panels do not have any role that would make them subject to the Federal Advisory Committee Act (FACA). They are used as information sources only, as sources for values elicitation (regarding weights and related matters), and so are decidedly not "advisory." They will never be asked for advice concerning the alternative solutions or the overall decision process.

Three Types of Evaluation Measures

For the Bay-Delta Program, three types of evaluation measures will be incorporated into the evaluation of solution alternatives. Most of the discussion to date has focused on the physical performance measures for the alternatives (improvement of conditions in the Bay-Delta system). These measures are developed as part of Step 2—Values Definition and described in item 1, below.

1. **Weighted sum of performance measures reflecting the impacts/performance of alternative solutions.** These scores will be developed by Values Definition Tasks 1 through 4. They represent a fairly mechanistic scoring, with a resulting ranking of alternatives according to each of perhaps several stakeholder perspectives, each perspective represented by a set of weights. This evaluation could include the pricing out of externalities earlier in this appendix, and so be the basis for solutions that involve such things as direct or indirect externality pricing for exports.
2. **Measures of process and institutional desirability.** In the course of preparing for the July 11 dry run and in subsequent analyses, several measures of desirability of alternative solutions have been defined that go beyond measures of impact/performance. The current version of the objectives hierarchy would include the objectives summarized immediately below. The first set of measures, under "Effective Solution," covers the impact/performance measures developed in Values Definition Tasks 1 through 3. The remaining sets of measures represent important broader considerations. While the measures in the left column below are all amenable to a weighted-sum evaluation, a less formal method of evaluation for the measures in the right column may be appropriate.

Effective Solution
(direct impacts/performance of the solution alternatives)
Ecosystem Quality
Water Quality
Water Supply
System Vulnerability

Sensitive Solution
(impacts incurred in the course of implementing a solution)
Socioeconomic
Environmental
Recreation
Hydropower
Navigation
Flood Control
Water Quality
Water Supply
Drainage

Responsive Process
Open
Accessible
Responsive
Collaborative
Balanced
Timely
Systematic
Understandable
Defensible
Reproducible

Achievable Solution
Affordable
Equitable
Implementable
Durable

Measures of overall desirability. The third type of evaluation measure tackles the problem of providing guidance to improve a solution alternative, and to selecting solution alternatives for the short list. Given that there will be multiple rankings of the solution alternatives, one ranking for each stakeholder perspective, the following measures can be used to identify overall desirability of an alternative. While the two types of measures listed above measure the performance of (score) each solution alternative separately for each stakeholder group, these measures show how an alternative performs across stakeholder groups. These measures are listed last because they are the least structured of the measure types. However, in one of their roles they are the first type of measure to be used, since they are used as guidance for assembling actions into preliminary solution alternatives in Step 3—Alternative Creation. Possible measures and considerations include:

- Breadth of Support—Evaluate a solution based on how many stakeholder perspectives rank it at the top, or rank it in the top two, or rank it in the top three.
- Simulated Majority Rule—Evaluate a solution based on how it would perform in a series of majority-rule elections among the solutions, assuming that each stakeholder would vote for the solution that scores highest on the weighted sum based on that stakeholder's weights.

- Simulated Bargaining—There are some mathematical ways to combine the scores across stakeholder groups in a way that can simulate the outcome of a fair bargaining process among the groups.
- Maximize Equity—There are two basically different concepts of equity to be considered. One is equity in outcome, i.e., how equally are the stakeholders' objectives met. This type of equity may involve very different changes from the status quo across stakeholder groups. Another concept of equity is equity in change from the status quo. This may involve very different qualities of outcome across stakeholder groups. Either case involves important fundamental difficulties in comparing strengths of preference across stakeholder groups. However, the concept of equity may be so fundamental to the evaluation of solutions that equity measures need to be included.
- Maximize Efficiency—This measure would maximize the sum of welfare across stakeholder groups, regardless of how evenly that welfare is distributed. Again, this measure involves important fundamental difficulties in comparing strengths of preference across stakeholder groups.
- Maximize Balance Among Objectives—This concept measures to what degree a solution addresses all objectives equally. It does not reflect stakeholder welfare, but it can still be used as a rationale for refining and selecting solutions.
- Balanced Restoration of Resources—This concept measures to what degree a solution improves conditions in each resource area as a proportion to the resource area's decline from historic levels. Again, this concept does not reflect stakeholder welfare, but it can still be used as a rationale for refining and selecting solutions.

Constructing the Evaluation Model

The most effective form of the evaluation model depends on the results of Step 4, Tasks 1 through 7. The model would probably take the form of a spreadsheet model, in Excel, with an overlay software model called DPL. The advantages of this approach include:

- Spreadsheet models are fairly understandable to people other than programmers. Partly for this reason, people tend to "trust" a spreadsheet model more than they do "black box " type models.
- Anyone with basic computer resources can run an Excel spreadsheet. A large workstation or minicomputer is not required. Therefore the Program can practice "democracy of analysis," as discussed in the Special Considerations section.
- The DPL overlay provides a clear graphical description of the relationships presented in the spreadsheet, using influence-diagram graphics.

- The DPL overlay allows the Program to run very efficient monte carlo - type runs of the spreadsheet (repeated runs of the spreadsheet with each run representing one way the system could perform). DPL can provide clear, simple probability-distribution graphics, so large uncertainties can be handled in an efficient way.
- The DPL overlay contains powerful sensitivity analysis features, with clear graphics, which allows the Program to intelligently develop the model in an iterative way.
- Once the model is relatively stable, DPL can automatically "lift" the Excel code into very fast, efficient C code, allowing the Program to make large monte carlo runs on desktop computers in a reasonable time.

Alternatively, the Program may determine that at least parts of the model can be coded in Extend, a versatile dynamic modeling environment with a clear graphical interface. However, it is probably not powerful enough to handle the spreadsheet-like level of complexity at which the Program will be operating.

The schedule and budget for Phase 1 do not allow the Program to develop a complete, maximum-analysis, actions-to-outcomes evaluation model. Even if a complete, maximum-analysis model is to be developed, the Program should begin this task the same way, with the first versions of the model coded in Excel and DPL because they allow for simple construction of the model which then provides guidance for how best to build the more complete model.